Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor

Joshua M. Gendron, José L. Pruneda-Paz, Colleen J. Doherty, Andrew M. Gross, S. Earl Kang, and Steve A. Kay[1]

Section of Cell and Developmental Biology, Division of Biological Sciences, University of California at San Diego, La Jolla, CA 92093-0130

Contributed by Steve A. Kay, January 9, 2012 (sent for review January 5, 2012)

The first described feedback loop of the *Arabidopsis* circadian clock is based on reciprocal regulation between TIMING OF CAB EXPRESSION 1 (TOC1) and CIRCADIAN CLOCK-ASSOCIATED 1 (CCA1)/LATE ELONGATED HYPOCOTYL (LHY). CCA1 and LHY are Myb transcription factors that bind directly to the *TOC1* promoter to negatively regulate its expression. Conversely, the activity of TOC1 has remained less well characterized. Genetic data support that TOC1 is necessary for the reactivation of *CCA1/LHY*, but there is little description of its biochemical function. Here we show that TOC1 occupies specific genomic regions in the *CCA1* and *LHY* promoters. Purified TOC1 binds directly to DNA through its CCT domain, which is similar to known DNA-binding domains. Chemical induction and transient overexpression of *TOC1* in *Arabidopsis* seedlings cause repression of *CCA1/LHY* expression, demonstrating that TOC1 can repress direct targets, and mutation or deletion of the CCT domain prevents this repression showing that DNA-binding is necessary for TOC1 action. Furthermore, we use the Gal4/UAS system in *Arabidopsis* to show that TOC1 acts as a general transcriptional repressor, and that repression activity is in the pseudoreceiver domain of the protein. To identify the genes regulated by TOC1 on a genomic scale, we couple *TOC1* chemical induction with microarray analysis and identify previously unexplored potential TOC1 targets and output pathways. Taken together, these results define a biochemical action for the core clock protein TOC1 and refine our perspective on how plant clocks function.

transcriptional regulation | transcriptional potential | gene networks | circadian outputs | diurnal

Most organisms that experience day/night cycles have a circadian clock that phases cellular processes and behavior to specific times of day while also anticipating daily diurnal changes to confer a fitness advantage (1). The basic molecular architecture of most clocks consists of negative-feedback loops where positive and negative components control each other's expression to generate oscillations with an approximate 24-h period (1).

At the core of the *Arabidopsis* clock, CIRCADIAN CLOCK-ASSOCIATED 1 (CCA1) (2) and LATE ELONGATED HYPOCOTYL (LHY) (3) are morning-expressed Myb transcription factors that directly bind the evening element (4, 5) in the promoters of evening-expressed genes and act as transcriptional repressors. CCA1 and LHY have redundant functions (6, 7), are often coexpressed, and make up the negative arm of the first described feedback loop in *Arabidopsis* by binding the promoter of *TIMING OF CAB EXPRESSION 1* (*TOC1*) (4). Genetically, TOC1 is the positive component of this feedback loop because *CCA1/LHY* morning reactivation is dependent on TOC1 (4), yet *TOC1* overexpression also results in lower *CCA1/LHY* expression (8, 9), confusing our understanding of TOC1's role in the core clock feedback loop. TOC1 is an evening-expressed protein (10) that is part of a five-member family called the PSEUDO-RESPONSE REGULATORS (PRRs) that are expressed in succession from the morning to the night in the order PRR9, PRR7, PRR5, PRR3, and then TOC1 (11). Sequence similarity of these proteins occurs in two domains, the pseudoreceiver (PR) domain in the N terminus and the CCT [CONSTANS (CO), CO-like, TOC1] domain in the C terminus (12). The PR domain allows homo- and hetero-oligomerization between the PRRs and bridges interactions of the PRRs and other proteins (13–15). The CCT domain of CO, a direct transcriptional activator of *FLOWERING LOCUS T*, was recently shown to bind DNA (16). This function has not been reported for the TOC1/PRR CCT domains, although TOC1 and the PRRs occupy regions of the *CCA1/LHY* promoter in ChIP experiments (17–19). It was shown for TOC1 and hypothesized for the PRRs that this association with chromatin could be bridged by other DNA-binding factors, such as CCA1 Hiking Expedition (CHE) (17, 19). The intermediate region (IR) between the PR and CCT domains of PRR5, -7, and -9 contain a transcriptional-repression motif (19), however this motif is not present in the IR of TOC1.

In this work we show that full-length TOC1 binds to a *cis*-element we have termed T1ME (TOC1 morning element), and that the CCT domain from TOC1 or the PRRs is sufficient to bind directly to DNA. Furthermore, our results show that TOC1 acts as a general transcriptional repressor that negatively regulates *CCA1/LHY* and a suite of genomic targets involved in critical plant functions. The repression activity lies in the PR domain of TOC1 but relies on the presence of a functional CCT domain to negatively regulate its targets. These results provide biochemical and molecular evidence that TOC1 is part of a previously unknown family of DNA-binding transcriptional regulators. It also completes the first described feedback loop in the *Arabidopsis* clock by linking *CCA1/LHY* expression to direct DNA-binding and transcriptional regulation by TOC1.

Results

TOC1 Occupies Specific Regulatory Regions of *CCA1* and *LHY*. Previously it was shown that TOC1 occupies a region of the *CCA1* promoter that contains a TCP binding site (TBS) (17), presumably through interaction with the TCP transcription factor CHE. To explore whether TOC1 associates with other regions of the *CCA1* genomic region, we performed ChIP using the previously described *TOC1* minigene (*TOC1p::TOC1:YFP*) transgenic line (15). Seedlings were grown under 12-h light:12-h dark (LD) cycles and cross-linked at zeitgeber time (ZT) 14 when the peak of TOC1 protein accumulation occurs (15). Enrichment of *CCA1* genomic regions after immunoprecipitation with an anti-GFP antibody was determined by quantitative real-time PCR (qRT-PCR) using primers that cover the *CCA1* promoter and 5′ UTR regions (Fig. 1*A*). Our results indicate that TOC1 occupies three regions of the *CCA1* promoter and *CCA1* coding sequence

Author contributions: J.M.G., J.L.P.-P., and S.A.K. designed research; J.M.G., J.L.P.-P., C.J.D., and S.E.K. performed research; J.M.G., J.L.P.-P., A.M.G., and S.E.K. contributed new reagents/analytic tools; J.M.G., J.L.P.-P., C.J.D., A.M.G., and S.A.K. analyzed data; and J.M.G. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE34897).

[1]To whom correspondence should be addressed. E-mail: skay@ucsd.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1200355109/-/DCSupplemental.

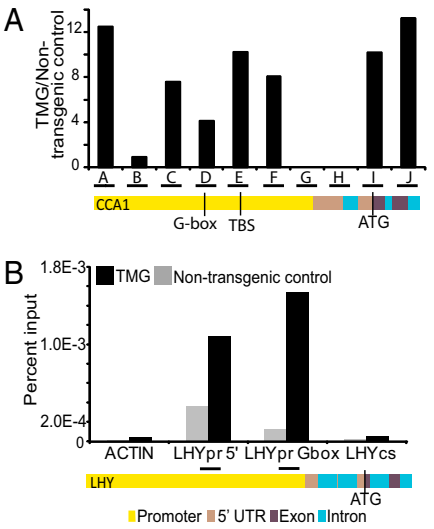

Fig. 1. TOC1 occupies multiple regions of the *CCA1* and *LHY* promoters. (*A*) ChIP–qRT-PCR of *TOC1* minigene (TMG) at the *CCA1* promoter. Values are relative to a nontransgenic control. (*B*) ChIP–qRT-PCR of three areas of the *LHY* genomic region. Values of the *TOC1* minigene and the nontransgenic control are calculated as percentage of the input. Relative placement of the amplified regions is marked by black lines. (Regions: A = −866/−786, B = −668/−589, C = −472/−365, D = −364/−274, E = −264/−171, F = −172/−85, G = −58/+69, H = +144/+235, I = +322/+413, J = +537/+622 base pairs from the transcriptional start site).

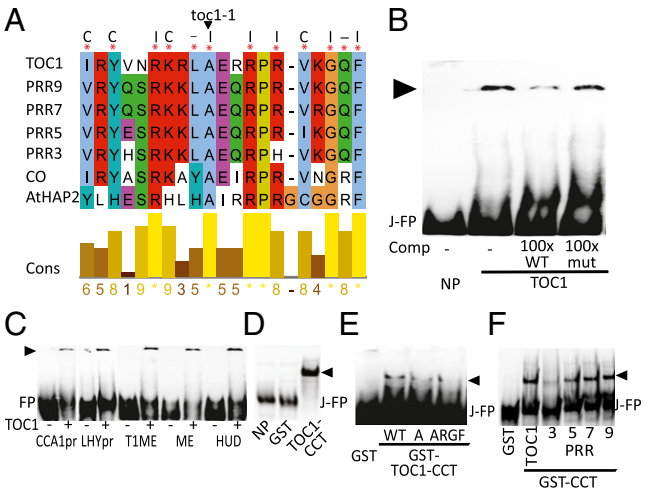

Fig. 2. TOC1 and the PRRs bind DNA directly through their conserved CCT domains. (*A*) ClustalW alignment of the C-terminal portions of the CCT domains of TOC1, PRR3, -5, -7, -9, and CO are aligned with the DNA-binding domain of AtHAP2. Stars mark amino acids that are necessary for DNA binding in the yeast HAP2 protein; "I" marks identical amino acids between all of the proteins (except Arginine 5, which is not conserved in PRR3); "C" marks conservation between TOC1/PRRs and CO; and "–" marks no conservation between TOC1/PRRs and CO or AtHAP2. The arrow points to the location of the *toc1-1* A to V mutation. Conservation (Cons) is plotted below. (*B*) EMSA assay with full-length TOC1 incubated with a probe designed to the J region (Fig. 1*A*) of the *CCA1* genomic region and either no competitor (-), 100-fold unlabeled wild-type competitor (100× WT), or 100-fold unlabeled mutant competitor (100× mut). NP, no protein. (*C*) EMSA with the full-Length TOC1 protein (FL) or no protein (NP) incubated with probes from the *CCA1* promoter (CCA1pr), *LHY* promoter (LHYpr), or synthetic probes containing the T1ME, ME, or HUD. (*D*) EMSA (J probe) with GST-tagged CCT from TOC1 (TOC1-CCT), GST alone (GST), or no protein (NP). (*E*) EMSA (J probe) with either the wild-type CCT domain of TOC1 or CCT domain with A or ARGF mutations. (*F*) EMSA (J probe) with the CCT domains of TOC1, PRR3, -5, -7, and -9. FP, free probe and J, J region of the *CCA1* promoter. Arrowheads mark shifted band.

outside of the TBS region (E and F region) (Fig. 1*A*). Regions correspond as follows: A = −866/−786, B = −668/−589, C = −472/−365, D = −364/−274, E = −264/−171, F = −172/−85, G = −58/+69, H = +144/+235, I = +322/+413, or J = +537/+622 base pairs from the transcriptional start site. TOC1 is associated with a region upstream of the TBS that contains a G-Box (CACGTG) (C and D region), a region 5′ of the G-box region (A region), and a region near the beginning of the *CCA1* coding sequence (I and J regions) (Fig. 1*A*). It is possible that TOC1 is recruited to the site containing the G-box through its interaction with phytochrome interacting factor (PIFs) or phytochrome interacting factor-like (PILs) (20), but the I/J and A regions have no canonical *cis*-elements for transcription factors that are known to interact with TOC1 (G-box or TBS). This finding suggests that TOC1 either interacts with unknown DNA-binding factors that recruit it to these regions or that it can bind DNA directly. Previous work showed that the *LHY* promoter does not contain a TBS and is not bound by CHE in yeast one-hybrid assays (17). However, genetic evidence suggests that TOC1 also regulates the expression of *LHY* (4). To test whether TOC1 can associate with the *LHY* promoter, we used ChIP–qRT-PCR with primers designed to two regions of the *LHY* promoter (Fig. 1*B*), one containing a G-box and another without G-box or TBS that is located 5′ of the G-box in the *LHY* promoter (Fig. 1*B*). TOC1 associates at both of these locations in the *LHY* promoter. These experiments provide evidence that TOC1 can occupy the *LHY* promoter and indicates that TOC1 is directly associated to multiple regions in the promoters of *CCA1* and *LHY*.

TOC1 Binds DNA at the T1ME. To investigate whether TOC1 has DNA-binding activity, we examined the conserved regions of TOC1 and the PRRs and found, as expected, a high degree of similarity within the CCT domain (Fig. 2*A*). The N terminus of the CCT domains is more divergent than the C terminus (Fig. S1) and is known, in the HAP2 protein from yeast, to bridge interactions with other proteins (12, 21); the C-terminal portion of all CCT domains contain most of the amino acids necessary for DNA binding previously reported for CO and the DNA-binding protein AtHAP2 (Fig. 2*A*) (16, 21). To test whether TOC1 can bind DNA, purified full-length TOC1 protein was obtained using a baculovirus-insect cell purification system and used to perform EMSAs. As shown in Fig. 2*B*, TOC1 binds to a 60-bp probe from the J region of *CCA1* (Fig. 1*A*) that contains no TBS or G-box. The CO-response element was shown to be TGTG(N2-N3)ATG (16), but the probe contains only a TGTG motif, so we performed a competition EMSA with both wild-type and mutant (TGTG changed to AAAA) unlabeled J probes (Fig. 2*B*). Fig. 2*B* and Fig. S2 show that the wild-type probe competes with the labeled probe but the mutant probe does not, indicating that TOC1 binds to the TGTG motif. Interestingly, this TGTG sequence (underlined) that we termed T1ME is part of the previously defined morning element (ME) (CCACAC or GTGTGG) (22) and the hormone up-regulated at dawn motif (HUD) (CACATG or CATGTG) (23), which are enriched in the promoters of morning-phased genes. As shown in Fig. 2*C*, we performed EMSAs that confirmed that TOC1 binds to 60-bp synthetic probes containing the T1ME, ME, or HUD elements. It is known that the *LHY* and *CCA1* promoter regions can confer cycling (24), and the J region of the *CCA1* promoter is not necessary for the cyclic behavior of *CCA1* expression. To test whether TOC1 can bind directly to the promoters of *CCA1*/*LHY*, probes that contain the T1ME in the *CCA1* and *LHY* promoters were made. TOC1 is able to bind both probes (Fig. 2*C*), showing that TOC1 alone directly interacts with the promoter sequence

in vitro. To test whether the CCT of TOC1 is the domain that mediates direct DNA binding, the interaction between a purified recombinant GST-tagged TOC1-CCT (TOC1-CCT) and the J probe were tested by EMSA (Fig. 2*D*). The TOC1-CCT variant binds DNA and the GST and no-protein controls showed no DNA-binding activity. To determine if binding is specific to the CCT domain and whether conserved amino acids are necessary for DNA binding, two CCT-domain mutant variants were generated: a single alanine to valine (A mutation), which is found in the *toc1-1* semidominant mutant (6), and a quadruple alanine, arginine, glycine, and phenylalanine to valine, isoleucine, valine, and leucine (ARGF mutation), that has been shown to disrupt DNA binding in the HAP2 DNA-binding domain (Fig. 2*A*) (21). Both the A and ARGF mutant versions of the CCT domain lost affinity for the J probe (Fig. 2*E*), confirming that the CCT domain is sufficient for DNA binding and is dependent on the ARGF conserved amino acids.

The PRR proteins have been shown to be transcriptional repressors that associate with the *CCA1*/*LHY* promoters in ChIP experiments (19). Based on the results obtained with the TOC1-CCT domain, we investigated whether PRR-CCT domains encompass a similar DNA-binding activity. The four PRR CCT domains bound to the J probe, with PRR3 showing weak binding (Fig. 2*F*). This finding may coincide with a divergence in the PRR3 protein sequence at the fourth conserved arginine that is important for DNA binding in HAP2 (Fig. 2*A*). Taken together, these results indicate that TOC1 and the PRRs are a family of DNA-binding proteins and that the DNA-binding activity relies on the CCT domain shared by these proteins.

TOC1 Acts as a Transcriptional Repressor. Because a *toc1* loss-of-function mutant has low *CCA1*/*LHY* expression, it was proposed that TOC1 is an activator of *CCA1*/*LHY* (4), but conversely *TOC1* overexpression results in low *CCA1*/*LHY* expression (8, 9). To address this discrepancy and test whether TOC1 activates or represses *CCA1*/*LHY*, we used a published ethanol-inducible *TOC1* transgenic line (*ALC::TOC1*) (25). We treated both the transgenic line and wild-type with either a mock treatment or 2.5% ethanol for 3 h starting at ZT22 in LD and collected samples each hour. As shown in the upper left panel of Fig. 3*A*, *TOC1* is induced approximately sixfold by the ethanol treatment. At the same time points, *CCA1* and *LHY* are repressed up to twofold, suggesting that TOC1 acts as a repressor of these genes (Fig. 3*A*). We also performed the same experiments starting at ZT6 and ZT18 in constant light conditions (LL) and completed longer time courses (Fig. S3). *TOC1* induction represses *CCA1* and *LHY* expression under all cases except for the 1- and 7-h time points collected at ZT19 and ZT25 (Fig. S3 *B* and *C*).

To confirm that TOC1 represses *CCA1* expression, a full-length *TOC1* overexpression construct was cobombarded with a *CCA1::LUC*$^+$ reporter (*CCA1* promoter driving expression of firely luciferase) into *Arabidopsis* seedlings. Luciferase levels were threefold lower than the control when *TOC1* is overexpressed (Fig. 3*B* and Fig. S3*A*), confirming that TOC1 can repress *CCA1* expression. To test if this result is dependent on a functional CCT DNA-binding domain, we introduced the A and ARGF mutations (Fig. 2 *A* and *F*) into the CCT domain of the full-length *TOC1* and cobombarded these with the *CCA1::LUC*$^+$ reporter. Without a functional DNA-binding domain, TOC1 is not able to repress *CCA1* at both ZT1 and ZT13 (Fig. 3*B* and Fig. S4*A*). The expression levels of wild-type and mutant *TOC1* were quantified by qRT-PCR and confirmed that the effects were not because of differences in expression level (Fig. S4*B*).

To test whether TOC1 acts as a general transcriptional repressor, we used the Gal4-LexA/UAS system, which tests proteins for positive or negative transcriptional potential (26). TOC1 represses the activity of VP16 and did not activate the reporter on its own (Fig. 3*C*), showing that TOC1 is a transcriptional repressor. PRR5, -7, and -9 contain a repression motif in the IR of the protein that is not present in TOC1, so we also tested the three domains of TOC1 (PR, IR, and CCT) (Fig. 3*C*, cartoon under plot) for transcriptional potential. The results in Fig. 3*C* show that the TOC1 PR domain contains the transcriptional repression potential, supporting the idea that TOC1 does not contain the transcriptional repression motif in the IR as was seen for PRR5, -7, and -9 (19).

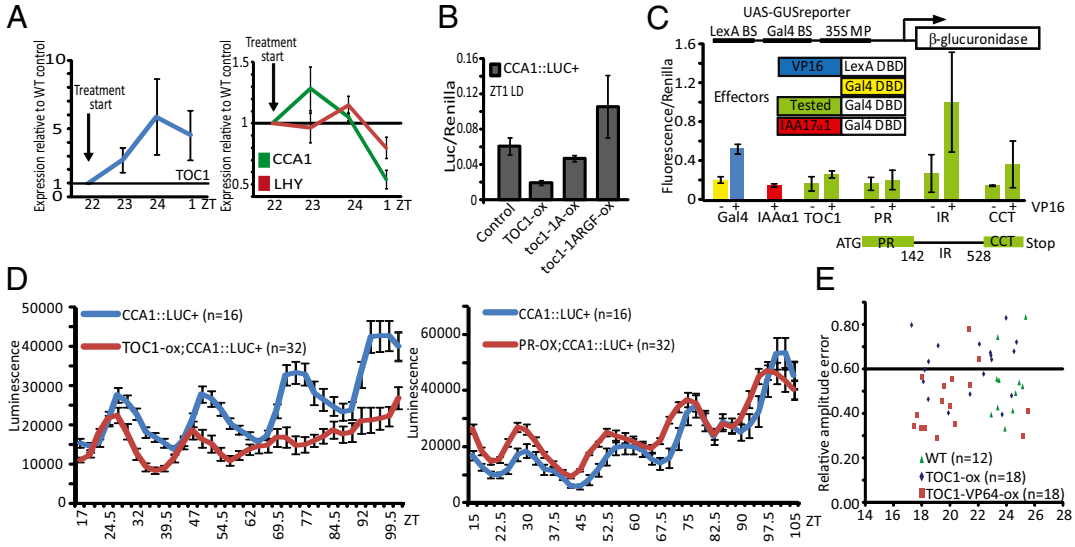

Fig. 3. TOC1 acts as a general transcriptional repressor. (*A*) qRT-PCR of *TOC1* (*Left*), *CCA1* (*Right*, green), *LHY* (*Right*, red) from *ALC::TOC1* induction experiment. (*B*) Luminescence measurements from seedlings overexpressing *TOC1*, *toc1-1A*, or *toc1-1ARGF* with the *CCA1::LUC*$^+$ reporter collected at ZT1 in LD. Error bars represent SD. (*C*) The Gal4/UAS transcriptional potential assay design is illustrated in the cartoon above the graph. The cartoon below the graph represents the exact TOC1 protein domains used in the assay. (*D*) Average luminescence of 16 *CCA1::LUC*$^+$ transgenic lines and 32 lines of *CCA1::LUC*$^+$ transformed with *TOC1-ox* (*Left*) or *PR-ox* (*Right*) in the T1 generation. (*E*) Plot of period against the RAE for *CCA1::LUC*$^+$ (green triangle) or *CCA1::LUC*$^+$ transformed with *TOC1-ox* (blue diamond) or *TOC1-VP64-ox* (red square) in the T1 generation. The line marks RAE of 0.6, which is a high cutoff for a gene to be considered cycling. Error bars represent SD (*A*–*C*) or SE (*D*).

To determine whether the PR domain can repress *CCA1* expression in the absence of the CCT and IR, we overexpressed the PR domain in a transgenic line carrying a *CCA1::LUC*$^+$ reporter construct and monitored oscillations in luciferase activity in LL (Fig. 3*D*). We averaged the luciferase activity of 32 T1 lines where either *TOC1* or *TOC1-PR* are overexpressed, and 16 *CCA1::LUC*$^+$ control seedlings. Overexpression of the full-length *TOC1* caused dampening of *CCA1* expression and arrhythmia after 2 d in LL (Fig. 3*D* and Fig. S5), as expected (8, 9). However, *TOC1-PR* overexpression caused a shortened period in luciferase activity (Fig. 3*D* and Fig. S5). This finding corresponds with previous findings that a nonfunctional, or in this case absent, CCT domain prevents TOC1 or the PR repression domain from negatively regulating *CCA1* expression (4, 6). Additionally, the short period caused by PR-constitutive expression is the same as seen in the *toc1-1* semidominant mutant and *toc1* loss-of-function mutants (4), suggesting that overexpression of the PR domain alone has the same effect on the clock as a loss-of-function mutant.

We next reversed the transcriptional polarity of TOC1 by making a translational fusion with the strong transcriptional activation domain, VP64. We overexpressed the full-length *TOC1* with (*TOC1-VP64-ox*) and without (*TOC1-ox*) the VP64 in a *CCA1::LUC*$^+$ genetic background and tracked luciferase activity in 18 T1 generation *TOC1-ox* or *TOC1-VP64-ox* lines and 12 *CCA1::LUC*$^+$ control lines. As was seen previously, most of the *TOC1* overexpression lines are arrhythmic [11 of 18 with a relative amplitude error (RAE) > 0.6] (Fig. 3*E*) or have a short period call when rhythmic (RAE < 0.6), whereas the *TOC1-VP64-ox* lines are mostly rhythmic (16 of 18 with RAE < 0.6) with a short period (Fig. 3*E*). This finding shows that reversing the transcriptional potential of TOC1 makes it behave like a loss-of-function mutant and confirms the importance of TOC1 repression activity to proper clock function.

TOC1 Controls a Suite of Clock-Regulated Genes. To identify potential targets and output pathways controlled by TOC1, we performed microarray analysis on RNA collected from *ALC::TOC1* and wild-type *Arabidopsis* mock-treated or ethanol-treated for 3 h and collected at ZT1 in either LD or LL. We treated all samples with 20 μM MG132 to block proteasome function and increase TOC1 protein levels after induction. Differentially expressed genes specific to *ALC::TOC1* and ethanol treatment and not wild-type or mock-treated were identified using LIMMA (27). This approach controls for the general effects of ethanol on *Arabidopsis* and for any gene-expression changes caused by the *ALC::TOC1* transgene before treatment.

In total, *TOC1* induction caused misexpression of 2,566 genes, with 1,254 genes up-regulated and 1,312 genes down-regulated (Fig. 4*A*). Of the up-regulated genes, 570 were misregulated in LD and 860 were misregulated in LL, with an overlap of 176 genes. Of the down-regulated genes, 580 were misregulated in LD and 876 were misregulated in LL. There was only one gene that overlapped between the genes that were down-regulated in LD and those that were up-regulated in LL, and no genes were shared between the genes that were down-regulated in LL and up-regulated in LD, showing that there was good agreement between the up-regulated and down-regulated genes from all datasets. The groups of genes were then compared with genes that were called cycling from four different microarray datasets, two collected from plants growing under LD conditions [LDHH_SM (28) and LDHH_ST (29)] and two collected from plants growing under LL conditions [LL12_LDHH (5) and LL23_LDHH (30)]. The genes that change by *TOC1* induction in LD and LL had more cycling genes in all four of the tested datasets than would be expected by chance, showing that TOC1 regulates expression of genes with cyclic behavior (Fig. 4*B*).

A permutation-based *z*-score was calculated to compare the phase distribution of the cycling genes that change by *TOC1*

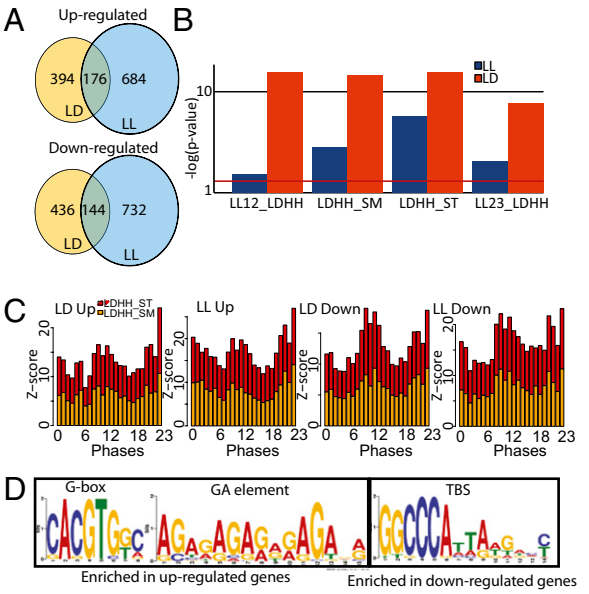

Fig. 4. TOC1 regulates a suite of clock genes and clock output genes. (*A*) Venn diagrams of total number of genes either up- or down-regulated by ethanol induction of *TOC1*. (*B*) Enrichment plots for the four experimental datasets used to determine cycling genes. LL (blue) and LD (red) refer to the growth conditions for the *TOC1*-induction experiment. The data for cycling genes were gathered from four microarray datasets listed on the *x* axis. Significance is plotted as the negative log of the *P* value obtained from Fisher's exact test. The red line marks significance threshold *P* value of 0.05 (1.30). (*C*) A *z*-score by phase for the enrichment of cycling genes that are regulated by *TOC1* induction. Red is from the LDHH_ST dataset and orange is from the LDHH_SM dataset. (*D*) WebLogo representation of overrepresented *cis*-elements in the promoters of the genes that are up- (CACGTGKM and AGARRGARRRAGADR) or down- (GGCCCA) regulated in both LD and LL.

induction (Fig. 4*C*). The genes up-regulated and down-regulated by *TOC1* induction in LDHH entrainment show a peak of enrichment at the morning and evening light transitions for both the LL and LD (Fig. 4*C*) induction conditions, with the highest peaks before dawn or subjective dawn. There is also a peak at dusk (LD) or subjective dusk (LL). The peaks seen during the dawn transition may represent the direct targets of TOC1 because those targets are expected to be phased to the morning, and the peak in the middle of the day may represent downstream targets of the morning-expressed transcription factors that are regulated by TOC1. The LL12 dataset shows a similar pattern of peaks as LDHH (Fig. S6*A*), and the peaks in the LL23 dataset are shifted (Fig. S6*A*). When all datasets are combined (Fig. S6*B*), the peaks follow the same trend as the LDHH (Fig. 4*C*).

To look for enrichment of *cis*-elements we took 500 bp of the promoter sequence from the shared up- (176 from Fig. 4*A*) and down- (144 from Fig. 4*A*) regulated genes from our dataset. These sequences were analyzed by three different motif-finding software packages [MEME (31), WEEDER (32, 33), and the R package rGADEM (34)]. Three elements were found to be enriched in the promoters of these genes using this stringent approach. In the down-regulated genes, a variant of the TBS (GGGCCAT) (35) was found (Fig. 4*D*). In the up-regulated genes a G-box variant (CACGTGKM) and a plant-specific GA motif (AGARRGARRRAGADR) are enriched. The GA motif is the *cis*-element for a family of DNA-binding proteins called BASIC PENTACYSTEINES (BPCs) (36). Interestingly, three of the four regions that we show to be occupied by TOC1 in the *CCA1* promoter contain the enriched elements. The E/F regions contain a TBS, the C/D regions contain a G-box, the A region contains a site similar to the AGARRGARRRAGADR element

(AAGGAGGAAGAAG or ARRGARRRAGADR) (Fig. 1*A*), and we have shown direct binding through the T1ME for the I/J region. These results suggest that there are complex transcriptional mechanisms underlying TOC1 regulation of gene expression. The T1ME was not found to be enriched in this dataset, but this may be expected because we have not defined the full *cis*-element sequence using an unbiased method, such as protein-binding microarray or systematic evolution of ligands by exponential enrichment. We have listed all genes that contain a T1ME, HUD, or ME in Dataset S1, and despite not being enriched in the dataset, the presence of these elements in a promoter may indicate direct binding by TOC1.

To find clock output pathways we examined genes differentially regulated by *TOC1* induction at least 1.8-fold for gene ontology (GO) enrichment in biological processes. Flow charts of these results are presented to visualize the hierarchy of GO categorization (Figs. S7 and S8). The down-regulated genes in LL were enriched in two main categories: response to stimulus, which includes stress, and biotic stimulus and metabolic process (Fig. S7). This finding coincides with the circadian clock being a major regulator of daily metabolic processes (37) and it is likely that TOC1 will be a hub that regulates genes involved in metabolism directly. Up-regulated genes in the LD experiment showed enrichment of categories of abiotic stimuli—such as light, temperature, and oxidative stress—all categories that coincide with the dark-to-light transition in the morning. This finding is in contrast to the down-regulated genes in LD, which show an enrichment for genes involved in biotic responses, mainly defense responses, which has recently been shown to be circadian-regulated (38, 39). These experiments show that TOC1 is involved in the regulation of circadian-regulated pathways that are critical to plant function.

Discussion

Since the identification of TOC1, extensive genetic analysis has placed TOC1 as the positive arm of the first described feedback loop in *Arabidopsis* (4, 9, 40). Using protein alignment, we show that TOC1/PRRs have a domain similar to known DNA-binding domains (12, 16, 21), and EMSAs show that the full-length TOC1 interacts with DNA through its CCT domain. TOC1 can bind a site [T1ME (TGTG)] that is part of the CO response element [TGTG(N2-N3)ATG] (16) and also binds the ME (22) and HUD (23) motifs (Fig. 2 *B* and *C*). We have also identified a set of three *cis*-elements that are enriched in genes regulated by TOC1 (Fig. 4*D*). Two of these sites are known *cis*-elements (G-box and TBS) that are binding sites for TOC1 interactors (PIFs/PILs and CHE), and the third is similar to a GA motif (AGARRGARRRAGADR), which is recognized by the BPC transcription factors. From our ChIP experiment, all three sites are present in regions of the *CCA1* promoter that TOC1 occupies (Fig. 1*A*). The enrichment of these elements in TOC1-controlled genes and the presence of TOC1 near these sites in vivo suggests that TOC1 recruitment to these elements may be important for the regulation of TOC1 targets. In combination with our data that a functional TOC1 DNA-binding domain is necessary for TOC1 function (Fig. 3 *B* and *D*), it is possible that TOC1 may heterodimerize with various transcription factors and directly bind DNA at or near their respective *cis*-element. It is easy to imagine that TOC1 is not only being recruited as a cofactor but that it participates in DNA binding with its partners to provide specificity for different *cis*-elements (41). Thus, as a monomer or homodimer TOC1 binds to the T1ME or similar site, but as a heterodimer TOC1 may associate with different *cis*-elements dependent on its protein interaction context. This process could be a mechanism for altering the phase of TOC1 targets by selective interaction with proteins, the temporal expression of which overlaps with TOC1.

Genetic mutation and transgenic overexpression of *TOC1* has provided a confusing picture of the role of TOC1 in the core loop of the clock (8, 9). It is clear from our results that TOC1 functions as a general transcriptional repressor (Fig. 3). We have not ruled out the possibility that TOC1 may activate gene expression under specific circumstances, but our evidence shows that TOC1 represses *CCA1*/*LHY* expression (Fig. 3 *A* and *B*) and the activity of a general transcriptional activator, VP16 (Fig. 3*C*). Additionally, mutation or deletion of the CCT shows that it is necessary for this repression activity, demonstrating that DNA-binding is integral to its function as a repressor (Fig. 3 *B* and *D*). Supporting this finding, mutations in the DNA-binding domain of NF-YA (a mammalian ortholog of HAP2) behave as dominant-negative alleles (42), which explains why the *toc1-1* mutation, as well as the overexpression of the PR domain, act in a dominant-negative manner, causing the short-period phenotype of *toc1* loss-of-function mutants but not lower *CCA1*/*LHY* expression levels. These data do not dispute the genetic data that TOC1 is necessary for *CCA1*/*LHY* reactivation; thus, it must still be considered the positive arm of the feedback loop genetically, maintaining the standard feedback structure of eukaryotic clocks. How TOC1 can both repress *CCA1*/*LHY* but also be necessary for their increase in expression in the morning is a critical question to answer for the future. Mathematical modeling and synthetic biology have shown that a transcriptional network of repressors can have cyclic behavior when paired with a constitutive transcriptional activator and delay in the feedback (43). This finding is likely to be the case in *Arabidopsis*, because a *toc1;prr9;prr7;prr5* quadruple mutant has high arrhythmic *CCA1* expression (40).

The PRRs were shown to have a repression motif in the IR (19), and TOC1 has a repression motif in the PR domain. It is possible that the TOC1 PR domain bridges interaction with proteins that contain a repressor motif, such as the PRRs (19) or CHE (17), and that TOC1 does not have inherent repressor activity. This theory would imply that TOC1 could activate transcription when complexed with a transcriptional activator. Recently, the mammalian clock component LITTLE IMAGINAL DISCS (LID) was shown to alternate between repressor and activator activity depending on its interaction partners (44). Further oligomerization studies will enhance our understanding of how protein interaction affects the transcriptional potential of TOC1.

As a DNA-binding transcription factor, it is likely that TOC1 does not only regulate genes directly involved in clock function but also genes involved in output processes. We found that TOC1 regulates a suite of genes that exhibit cycling behavior with enrichment at dawn and dusk (Fig. 4 *B* and *C*). These genes fall into biological categories, such as metabolism, biotic stress, and abiotic stress, and have *cis*-elements that are known to be enriched in clock-regulated genes (Fig. 4*D*, and Figs. S7 and S8). These output pathways are early clues to understanding the TOC1 transcriptional network.

The results presented in this article show that the critical clock component, TOC1, is a transcriptional regulator that binds DNA, and further defines TOC1 transcriptional networks that control growth and development.

Materials and Methods

Plant Materials and Growth. *Arabidopsis* ecotype Colombia (Col-0) was used for all experiments. Growth conditions and transgenic lines are described in *SI Materials and Methods*.

Chromatin Immunoprecipitation. ChIP was done as described previously (17). Primer details are described in *SI Materials and Methods*.

Electrophoretic Mobility Shift Assays. EMSAs were done with standard conditions provided in the Lightshift Chemiluminescent EMSA kit (Thermo

Scientific). Protein purification conditions, EMSA buffers, and probe sequences are provided in *SI Materials and Methods*.

Ethanol Induction Followed by qRT-PCR or Microarray. An established *TOC1* inducible transgenic line (*ALC::TOC1*) (25) and Col-0 (wild-type) were used in all experiments. Growth and induction conditions, qRT-PCR primers, and microarray hybridization are described in *SI Materials and Methods*.

Gal4/UAS System. The Gal4/UAS system has been described previously (26). Experimental details are provided in *SI Materials and Methods*.

Microarray Analysis. Details of microarray analysis are described in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank B. Chow for critical reading of the manuscript; the S.A.K. laboratory members and S. B. for fruitful discussion; A. L. Quiroz, C. Valdivia, N. Navarro, and G. Cachu for technical support; and A. Lestelle for administrative support. This work is supported in part by National Institutes of Health Grants R01GM056006 and R01GM67837 (to S.A.K.); the Ruth L. Kirchstein National Research Service Award F32GM087114 (to J.M.G.); and National Science Foundation Postdoctoal Research Fellowship in Biology 0906055 (to C.J.D.).

1. Bell-Pedersen D, et al. (2005) Circadian rhythms from multiple oscillators: Lessons from diverse organisms. *Nat Rev Genet* 6:544–556.
2. Wang ZY, Tobin EM (1998) Constitutive expression of the *CIRCADIAN CLOCK ASSOCIATED 1* (*CCA1*) gene disrupts circadian rhythms and suppresses its own expression. *Cell* 93:1207–1217.
3. Schaffer R, et al. (1998) The late elongated hypocotyl mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering. *Cell* 93: 1219–1229.
4. Alabadí D, et al. (2001) Reciprocal regulation between TOC1 and LHY/CCA1 within the *Arabidopsis* circadian clock. *Science* 293:880–883.
5. Harmer SL, et al. (2000) Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science* 290:2110–2113.
6. Alabadí D, Yanovsky MJ, Más P, Harmer SL, Kay SA (2002) Critical role for CCA1 and LHY in maintaining circadian rhythmicity in *Arabidopsis*. *Curr Biol* 12:757–761.
7. Mizoguchi T, et al. (2002) LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in *Arabidopsis*. *Dev Cell* 2:629–641.
8. Makino S, Matsushika A, Kojima M, Yamashino T, Mizuno T (2002) The APRR1/TOC1 quintet implicated in circadian rhythms of *Arabidopsis thaliana*: I. Characterization with APRR1-overexpressing plants. *Plant Cell Physiol* 43:58–69.
9. Más P, Alabadí D, Yanovsky MJ, Oyama T, Kay SA (2003) Dual role of TOC1 in the control of circadian and photomorphogenic responses in *Arabidopsis*. *Plant Cell* 15: 223–236.
10. Strayer C, et al. (2000) Cloning of the *Arabidopsis* clock gene *TOC1*, an autoregulatory response regulator homolog. *Science* 289:768–771.
11. Matsushika A, Makino S, Kojima M, Mizuno T (2000) Circadian waves of expression of the APRR1/TOC1 family of pseudo-response regulators in *Arabidopsis thaliana*: Insight into the plant circadian clock. *Plant Cell Physiol* 41:1002–1012.
12. Wenkel S, et al. (2006) CONSTANS and the CCAAT box binding complex share a functionally important domain and interact to regulate flowering of *Arabidopsis*. *Plant Cell* 18:2971–2984.
13. Wang L, Fujiwara S, Somers DE (2010) PRR5 regulates phosphorylation, nuclear import and subnuclear localization of TOC1 in the *Arabidopsis* circadian clock. *EMBO J* 29:1903–1915.
14. Para A, et al. (2007) PRR3 Is a vascular regulator of TOC1 stability in the *Arabidopsis* circadian clock. *Plant Cell* 19:3462–3473.
15. Más P, Kim WY, Somers DE, Kay SA (2003) Targeted degradation of TOC1 by ZTL modulates circadian function in *Arabidopsis thaliana*. *Nature* 426:567–570.
16. Tiwari SB, et al. (2010) The flowering time regulator CONSTANS is recruited to the FLOWERING LOCUS T promoter via a unique *cis*-element. *New Phytol* 187:57–66.
17. Pruneda-Paz JL, Breton G, Para A, Kay SA (2009) A functional genomics approach reveals CHE as a component of the *Arabidopsis* circadian clock. *Science* 323: 1481–1485.
18. Legnaioli T, Cuevas J, Mas P (2009) TOC1 functions as a molecular switch connecting the circadian clock with plant responses to drought. *EMBO J* 28:3745–3757.
19. Nakamichi N, et al. (2010) PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the *Arabidopsis* circadian clock. *Plant Cell* 22:594–605.
20. Yamashino T, et al. (2003) A Link between circadian-controlled bHLH factors and the APRR1/TOC1 quintet in *Arabidopsis thaliana*. *Plant Cell Physiol* 44:619–629.
21. Xing Y, Fikes JD, Guarente L (1993) Mutations in yeast HAP2/HAP3 define a hybrid CCAAT box binding domain. *EMBO J* 12:4647–4655.
22. Michael TP, et al. (2008) Network discovery pipeline elucidates conserved time-of-day-specific *cis*-regulatory modules. *PLoS Genet* 4:e14.
23. Michael TP, et al. (2008) A morning-specific phytohormone gene expression program underlying rhythmic plant growth. *PLoS Biol* 6:e225–e225.
24. Salomé PA, Xie Q, McClung CR (2008) Circadian timekeeping during early *Arabidopsis* development. *Plant Physiol* 147:1110–1125.
25. Knowles SM, Lu SX, Tobin EM (2008) Testing time: Can ethanol-induced pulses of proposed oscillator components phase shift rhythms in *Arabidopsis*? *J Biol Rhythms* 23:463–471.
26. Tiwari SB, Wang XJ, Hagen G, Guilfoyle TJ (2001) AUX/IAA proteins are active repressors, and their stability and activity are modulated by auxin. *Plant Cell* 13: 2809–2822.
27. Smyth GK (2005) *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*, eds Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W (Springer, New York), pp 397–420.
28. Smith SM, et al. (2004) Diurnal changes in the transcriptome encoding enzymes of starch metabolism provide evidence for both transcriptional and posttranscriptional regulation of starch metabolism in *Arabidopsis* leaves. *Plant Physiol* 136:2687–2699.
29. Bläsing OE, et al. (2005) Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in *Arabidopsis*. *Plant Cell* 17: 3257–3281.
30. Edwards KD, et al. (2006) FLOWERING LOCUS C mediates natural variation in the high-temperature response of the *Arabidopsis* circadian clock. *Plant Cell* 18:639–650.
31. Bailey TL, Williams N, Misleh C, Li WW (2006) MEME: Discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res* 34:W369–W373.
32. Pavesi G, Mereghetti P, Mauri G, Pesole G (2004) Weeder web: Discovery of transcription factor binding sites in a set of sequences from co-regulated genes. *Nucleic Acids Res* 32:W199–W203.
33. Pavesi G, et al. (2006) MoD tools: Regulatory motif discovery in nucleotide sequences from co-regulated or homologous genes. *Nucleic Acids Res* 34:W566–W570.
34. Mercier E, et al. (2011) An integrated pipeline for the genome-wide analysis of transcription factor binding sites from ChIP-Seq. *PLoS ONE* 6:e16432.
35. Viola IL, Uberti Manassero NG, Ripoll R, Gonzalez DH (2011) The *Arabidopsis* class I TCP transcription factor AtTCP11 is a developmental regulator with distinct DNA-binding properties due to the presence of a threonine residue at position 15 of the TCP domain. *Biochem J* 435:143–155.
36. Wanke D, et al. (2011) Alanine zipper-like coiled-coil domains are necessary for homotypic dimerization of plant GAGA-factors in the nucleus and nucleolus. *PLoS ONE* 6:e16070.
37. Graf A, Smith AM (2011) Starch and the clock: The dark side of plant productivity. *Trends Plant Sci* 16:169–175.
38. Bhardwaj V, Meier S, Petersen LN, Ingle RA, Roden LC (2011) Defence responses of *Arabidopsis thaliana* to infection by *Pseudomonas syringae* are regulated by the circadian clock. *PLoS ONE* 6:e26968.
39. Wang W, et al. (2011) Timing of plant immune responses by a central circadian regulator. *Nature* 470:110–114.
40. Yamashino T, et al. (2008) Involvement of *Arabidopsis* clock-associated pseudo-response regulators in diurnal oscillations of gene expression in the presence of environmental time cues. *Plant Cell Physiol* 49:1839–1850.
41. Tucker SC, Wisdom R (1999) Site-specific heterodimerization by paired class homeodomain proteins mediates selective transcriptional responses. *J Biol Chem* 274: 32325–32332.
42. Mantovani R, et al. (1994) Dominant negative analogs of NF-YA. *J Biol Chem* 269: 20340–20346.
43. Stricker J, et al. (2008) A fast, robust and tunable synthetic gene oscillator. *Nature* 456:516–519.
44. DiTacchio L, et al. (2011) Histone lysine demethylase JARID1a activates CLOCK-BMAL1 and influences the circadian clock. *Science* 333:1881–1885.